Topics for Circumscription

Topics for Circumscription

by
Mary-Angela Papalaskaris and Alan Bundy

Acknowledgements

We are grateful to Mark Steedman, John Jones and Leon Sterling for their comments on a preliminary version of this paper. This research was supported by an Edinburgh University Postgraduate Studentship to M.A. Papalaskaris and SERC grant, number GR/C/20826.

1. Introduction

The power of human reasoning can, to a great extent, be attributed to people's ability to reason from incomplete information. People seem able to 'jump to conclusions' and make plausible conjectures which fill out the information they are given.

Circumscription is a rule of conjecture developed by McCarthy [McCarthy 80] that can be used to jump to certain conclusions from incomplete information. Intuitively circumscription states that the objects that can be shown to have a certain property **P** by reasoning from certain facts **A** are all the objects that satisfy **P**. In this sense it is related to the closed world assumption and to the negation as failure rule. In fact, circumscription gives roughly a more general form of the closed world assumption.

As with the rules of deduction, to which it is related, circumscription incorporates choice points. Deductive systems typically use heuristics to make these choices, and are prepared to remake them in a search for a desired conclusion. Unrestricted search is not appropriate when using circumscription; not just for combinatorial reasons, but because some of the choices lead to implausible conjectures. We must use the context to make the choices and avoid jumping to inappropriate conclusions. In this paper we propose a technique for using contextual information in this way.

2. Circumscription

The circumscription of sentence **A(P)** with respect to predicate **P** is the sentence schema:

$$\mathbf{A}(\Phi) \wedge \forall x(\Phi(x) \rightarrow P(x)) \rightarrow \forall x(P(x) \rightarrow \Phi(x)) \qquad \text{(i)}$$

Suppose we are given a sentence **A(P)** which we have reason to believe supplies all the relevant facts about objects with property **P**. Hence, any apparently weak property, **Φ**, for which **A** holds, is not really weaker but is equivalent to P. Circumscribing **A(P)** with respect to **P** allows us to jump to this conclusion.

(i) contains a predicate parameter **Φ** for which we may substitute an arbitrary predicate expression, thus we could view it as a second order sentence with a $\forall\Phi$ in front of it.

The decision as to which formulae to substitute for Φ, P and A in (i) all constitute choice points in the application of circumscription.

A circumscriptive inference, i.e., an inference which yields the conjecture **q** by circumscribing **P** in **A(P)**, is written $\mathbf{A} \vdash_{P} \mathbf{q}$.

Circumscriptive inference is non-monotonic. This means that when more facts are added to the theory it is possible that something ceases to be a theorem, i.e., the set of theorems does not increase monotonically with the set of axioms of the theory. The well known systems of mathematical logic have this monotonicity property. Thus circumscription can be used along with the rules of ordinary first-order logic, as a *rule of conjecture*, to give a system capable of non-monotonic reasoning.

A typical example of the use of circumscription (from [McCarthy 80]) involves circumscribing the predicate 'block' in a sentence of the form 'block(B)$\wedge$block(E)', i.e. substituting 'block' for P and 'P(B)$\wedge$P(E)' for A(P) in (i). This gives:

$$\Phi(B) \wedge \Phi(E) \wedge \forall x(\Phi(x) \rightarrow block(x)) \rightarrow \forall x(block(x) \rightarrow \Phi(x)) \quad \text{(ii)}$$

Choosing $\Phi(x) \equiv x{=}B \vee x{=}E$ and simplifying we have:
$\forall x(block(x) \rightarrow (x{=}B \vee x{=}E))$

Thus, the circumscriptive inference is:
$(block(B) \wedge block(E)) \vdash_{block} \forall x(block(x) \rightarrow (x{=}B \vee x{=}E))$

This can be paraphrased as "if B and E are blocks then they are the only blocks", which is an example of the closed world assumption. In some contexts this conjecture is intuitive and plausible. For instance, if we have asked "which of these are blocks" and been told B and E, then it is reasonable to jump to the conclusion that B and E are the only blocks. Note that obtaining this particular circumscriptive inference relies on choosing the substitutions for Φ, P and A appropriately. We will see that different choices lead to different conjectures.

The above example shows how circumscription can yield plausible conjectures in a neat and concise way. To see how circumscription can lead to implausible conjectures, consider the following example

Let P be: block
Let A(block) be: $block(B) \rightarrow block(E)$ (iii)

Circumscription of block in (iii):

$$(\Phi(B) \rightarrow \Phi(E)) \wedge \forall x(\Phi(x) \rightarrow block(x)) \rightarrow \forall x(block(x) \rightarrow \Phi(x)) \quad \text{(iv)}$$

1. Let $\Phi(x) \equiv x{=}C$. Then from (iv) and assuming $B \neq C$:

   $(block(B) \rightarrow block(E)) \vdash_{block} block(C) \rightarrow \forall x(block(x) \rightarrow x{=}C)$
   i.e., if C is a block then it is the only block.

2. Let $\Phi(x) \equiv$ **false**. Then from (iv):

$(\mathrm{block}(B) \rightarrow \mathrm{block}(E)) \vdash_{block} \forall x(\neg\mathrm{block}(x))$

i.e., there are no blocks.

This example shows the inappropriateness of unguided application of circumscription. The outcome that if C is a block then it is the only one, is a very unlikely one among humans. Perhaps a more desirable one is the one where implication is reversed, i.e., block(E)→block(B) [Johnson-Laird and Wason 77], but that is not possible under any substitution in this particular instance of the circumscription schema, i.e., for this particular choice of predicate to be circumscribed with respect to. As we see later (section 3), and as follows from [Reiter 82], it is possible to derive the reversal of the implication by circumscription, by choosing a different predicate to circumscribe with respect to.

3. Topics

The examples in the previous section show some of what are believed to be typical applications of circumscriptive inference. In some cases circumscription seems to model 'jumping to conclusions' very accurately: the inferences are very intuitive. In other cases, such as the last example, it seems less plausible. Moreover, even in the case of the first example, the inference that there are no blocks other than B and E is not always plausible. Consider, for instance, the same sentence, 'block(B)∧block(E)', as an answer to "what are B and E?". Clearly it is not valid to infer that B and E are the only blocks from the answer, in such a context.

In this paper we restrict our attention to the appropriate applications of circumscription to the answers given to questions and concentrate on formulating guidelines for the selection of the predicate to be circumscribed, based on contextual information contained in the question. We will find that one piece of contextual information is of decisive importance, namely the *topic* of the original question. The topic of the question is only part of the whole context of the question and answer session. The whole context can be arbitrarily large and complex, since it has to specify and define a situation along many dimensions. The topic for a statement is easier to pinpoint, for it concerns only one (or a few) of the dimensions. Thus the identification of the topic as the key component of the context considerably simplifies the process of guiding circumscription. Where the *same* answer is given to *different* questions, circumscription of the topic of the question in the answer will yield different conjectures – in each case the conjecture being plausible in the context of the question.

For example:

1. in light of the question, "what are B and E?" the topics would be $\lambda P[P(B)]$ and $\lambda P[P(E)]$*, i.e., the classifications for B and E.

2. On the other hand, a question such as "which of these are blocks?"

*Note that the question is ambiguous. One could interpret it also as meaning "what common properties do B and E have?" which yields $\lambda P[P(B) \wedge P(E)]$ as the topic. The same treatment applies to this case.

requires the answerer to mention all of the objects under consideration** that are blocks, so the topic in this case is $\lambda x[block(x)]$.

In these two different cases, circumscription ought to yield different conjectures. In (1), since the question is asking what B and E are, the answer can lead one to jump to the conclusion that all we know of B and E is that they are blocks, or else the answer would contain the additional properties. This is consistent with Grice's maxims of cooperative conversation [Grice 75], in particular the quantity maxim that states: "make your contribution as informative as is required". Similarly in (2), where the question now is as to which are blocks, the *same* answer would lead to the conclusion that B and E are the only blocks.*

The (answer) statement $block(B) \wedge block(E)$ can be rewritten as

$$\lambda P[\lambda x[P(x)](B)](block) \wedge \lambda P[\lambda x[P(x)](E)](block) \quad (v)$$

or

$$\lambda x[\lambda P[P(x)](block)](B) \wedge \lambda x[\lambda P[P(x)](block)](E) \quad (vi)$$

(both forms are equivalent)

The topics for (1) and (2) can also be put in the same form:

1. $\lambda P[\lambda x[P(x)](B)]$ and $\lambda P[\lambda x[P(x)](E)]$ (vii)
2. $\lambda P[\lambda x[P(x)]](block)$ (viii)

In this question/answer context the circumscription choices should be exercised as follows.

- From Grice's maxims we are justified in assuming that the answer to the query provides all the relevant facts about the topic of the question, i.e. it contains more information than it literally states. We can draw this extra information from it by circumscribing it. Thus the answer should be substituted for **A(P)**.

- The topic of the original question is what the answer provides more information about. Thus we should circumscribe with respect to the topic, i.e. we should substitute the topic for P.

**this restriction is of interest on its own right, but is not dealt with here

*It may be argued that these conclusions do not stem from conversational implicatures, but rather implicit meanings of the questions. "Which of these are blocks?", for example, can be taken to also require *all* the blocks in question. There are some cases however where such readings are not allowed. Consider, for example, the situation where you are in a hurry to build something out of a set of available objects and turn to your partner saying "Quickly, I need a block. Which of these are blocks?". In such a context it would be wrong to interpret his/her answer to provide *all* the blocks.

- On this analysis we are justified in making any substitution for Φ. In practice it is usually sufficient to substitute one of a number of simple predicates, such as F, where $F(x) \equiv$ **false** or a predicate already occurring in **A(P)**.

We illustrate this in detail beginning with the first case. For simplicity, $\lambda P[\lambda x[P(x)](B)]$ in (vii) is abbreviated to $\lambda P[B\text{-is}(P)]$, and $\lambda P[\lambda x[P(x)](E)]$ to $\lambda P[E\text{-is}(P)]$, although the same treatment can be carried out using the original λ-expressions. (v) then becomes:

$$B\text{-is}(block) \land E\text{-is}(block)$$

circumscribing B-is

$$\Phi(block) \land E\text{-is}(block) \land \forall x(\Phi(x) \rightarrow B\text{-is}(x)) \rightarrow \forall x(B\text{-is}(x) \rightarrow \Phi(x))$$

substituting $\Phi(x) \equiv x{=}block$, and simplifying:

$$\forall x(B\text{-is}(x) \rightarrow x{=}block) \qquad \text{(ix)}$$

and similarly, by circumscribing E-is(x):

$$\forall x(E\text{-is}(x) \rightarrow x{=}block) \qquad \text{(x)}$$

Thus in the first case, where the topic is the property of B and E, (ix) and (x) reflect our expectation for a conjecture that 'blockness' is the only (relevant) property of B and E.

For the second case (viii) is abbreviated to $\lambda x[block(x)]$ and (vi) becomes $block(B) \land block(E)$, which is the form used in the original example. Recall that by circumscriptive inference:

$$\forall x(block(x) \rightarrow (x{=}B \lor x{=}E))$$

So we see that the same procedure, but using a different topic, produces the result that B and E are the only blocks, which again reflects our expectation since the question was "which of these are blocks?".

The importance of using the notion of topic to guide circumscription is the uniformity of treatment: block(x) and B-is(P) are both instances of a more general λ-expression, $\lambda P[\lambda x[P(x)]](block)(B)$, which is just another way of writing the 'block(B)' that occurs in the answer sentence. Both instances can be used as topics for circumscription and produce different conjectures from the same sentence. The use of the topic of the question as the circumscribed predicate, P, is the same in each case, but the question asked leads to a different topic and, hence, a different conjecture.

Another example of the same treatment is illustrated by circumscription of the statement $block(B) \rightarrow block(E)$. We saw in the last example of the previous section that when the statement is taken to be about blocks in general, circumscription conjectures that there are no blocks. But suppose that the context makes it clear that the statement is intended as a definition of 'block(E)', then the topic will be 'block(E)' (i.e., $\lambda P[\lambda x[P(x)]](\text{block})(E)$) and circumscription will give a different conjecture, as shown below. By circumscription:

$$(\text{block}(B) \rightarrow \Phi) \wedge (\Phi \rightarrow \text{block}(E)) \rightarrow (\text{block}(E) \rightarrow \Phi)$$

substitution $\Phi \equiv \text{block}(B)$, and simplifying:

$$\text{block}(E) \rightarrow \text{block}(B)$$

This is a very desirable result because it happens to be a very frequent conclusion that people jump to, i.e., the reversal of an implication, and one that is difficult to capture with general methods [Johnson-Laird and Wason 77]. This example again illustrates the power that can be gained by viewing statements as very general λ-expressions and choosing the right level of instantiation - dictated by the topic - to circumscribe in.

The same result can also be obtained via predicate completion [Clark 78]. Reiter [Reiter 82] has shown that if T is a first order theory in clausal form, Horn in a predicate P, then P's completion axiom, $\forall x(P(x) \rightarrow A(x))$, is derivable by circumscription. So it is possible to view $\text{block}(B) \rightarrow \text{block}(E)$ as the theory, Horn in block(E) (here, again block(E) is taken as the predicate in question, with no free variables) and block(E)'s completion axiom is precisely the reversal of the implication. Note that the same underlying process is taking place here as with the use of topics, since a formula Horn in a predicate can be thought of as giving a definition for that predicate. Forcing $\text{block}(B) \rightarrow \text{block}(E)$ into the Horn clause framework before circumscribing has the effect of restricting the choice of predicate to be circumscribed, in much the same way as the topic does. The advantage of using the notion of topic is that the choice is made explicit and it provides a way of expressing and reasoning about it using contextual information.

4. Conclusion

Circumscription offers a powerful rule of conjecture which can be added to the traditional rules of deduction to model the human ability to jump to conclusions. However, unrestricted circumscription can yield inappropriate and implausible conclusions. Contextual information must be used to make the choices implicit in circumscriptive inference.

In the case of a simple question-answering situation, we have observed that the topic of the original question plays a decisive role in determining which of these choices is appropriate. A simple computational procedure is suggested by this observation. We are building a computer program which will implement this procedure and guide the process of circumscription.

We hope to extend the notion of topic to apply outside the realm of question answering. The last example suggests that this is possible, but more work needs to be done to determine how to extract topics from situations where there is no explicit question in hand.

References

[Clark 78] Clark, K. L.
Negation as failure.
In Gallaire, H. and Minker, J. (editors), *Logic and Data Bases*, pages 293-322. Plenum Press, 1978.

[Grice 75] Grice, H. P.
Logic and Conversation.
In Cole, P. and Morgan, J. L. (editors), *Syntax and Semantics: Speech Acts*, pages 41-58. Academic Press, 1975.

[Johnson-Laird and Wason 77]
Johnson-Laird, P. N. and Wason, P. C.
Readings in Cognitive Science: Thinking.
Cambridge University Press, 1977.

[McCarthy 80] McCarthy, J.
Circumscription - A Form of Non-Monotonic Reasoning.
Artificial Intelligence (13):27-39, 1980.

[Reiter 82] Reiter, R.
Circumscription implies predicate completion (sometimes).
In *Proceedings of AAAI-82*, pages 418-420. American Association for Artificial Intelligence, 1982.